

# Dual Averaging on Compactly-Supported Distributions And Application to No-Regret Learning on a Continuum

Walid Krichene

## Abstract

We consider an online learning problem on a continuum. A decision maker is given a compact feasible set  $S$ , and is faced with the following sequential problem: at iteration  $t$ , the decision maker chooses a distribution  $x^{(t)} \in \Delta(S)$ , then a loss function  $\ell^{(t)} : S \rightarrow \mathbb{R}_+$  is revealed, and the decision maker incurs expected loss  $\langle \ell^{(t)}, x^{(t)} \rangle = \mathbb{E}_{s \sim x^{(t)}} \ell^{(t)}(s)$ . We view the problem as an online convex optimization problem on the space  $\Delta(S)$  of Lebesgue-continuous distributions on  $S$ . We prove a general regret bound for the Dual Averaging method on  $L^2(S)$ , then prove that dual averaging with  $\omega$ -potentials (a class of strongly convex regularizers) achieves sublinear regret when  $S$  is uniformly fat (a condition weaker than convexity).

## 1 Introduction

We consider an online learning problem on a compact subset  $S \subset \mathbb{R}^n$ . At each iteration  $t \in \mathbb{N}$ , a decision maker chooses a distribution  $x^{(t)}$  on  $S$ , then, a loss function  $\ell^{(t)} : S \rightarrow \mathbb{R}_+$  is revealed, and the decision maker incurs loss  $\mathbb{E}_{s \sim x^{(t)}} [\ell^{(t)}(s)]$ . This is summarized in Problem 1.

---

**Problem 1** Online decision problem with Lipschitz losses on  $S$ .

---

- 1: **for**  $t \in \mathbb{N}$  **do**
  - 2:   Decision maker chooses distribution  $x^{(t)}$  over  $S$ .
  - 3:   A loss function  $\ell^{(t)} : S \rightarrow \mathbb{R}_+$  is revealed. We assume  $\ell^{(t)}$  is  $L$ -Lipschitz.
  - 4:   The decision maker incurs expected loss  $\mathbb{E}_{s \sim x^{(t)}} [\ell^{(t)}(s)]$
  - 5: **end for**
- 

The regret of the decision maker is defined as follows: for a given sequence of losses  $(\ell^{(t)})_{t \in \mathbb{N}}$ , and a corresponding sequence of decisions  $(x^{(t)})_{t \in \mathbb{N}}$ , the cumulative regret at time  $t$ , denoted by

$$R^{(t)} = \sum_{\tau=1}^t \mathbb{E}_{s \sim x^{(\tau)}} [\ell^{(\tau)}(s)] - \inf_{s \in S} \sum_{\tau=1}^t \ell^{(\tau)}(s)$$

compares the expected loss cumulated by the decision maker to the infimum of the cumulative loss function. In particular, we seek to design algorithms for which the regret grows sub-linearly in  $t$ , *for any sequence of losses* in a given class (the assumptions on the losses will later be made explicit).

This sequential decision problem has a long history which dates back to Hannan [14] and Blackwell [5], who formulated the problem in the context of repeated games. The notion of regret is closely related to the notion of consistent play (as defined by Hannan) and approachability (as defined by Blackwell). Beyond player dynamics in repeated games, online learning has many applications such as portfolio optimization [10, 6] and machine learning [12].

Regret minimization is essential in the design and analysis of online learning algorithms [9, 8], and the study of player dynamics in repeated games [15, 16, 23, 22]. In this article, we study the problem of designing sublinear regret algorithms under minimal assumptions on the feasible set  $S$  and the sequence of losses  $(\ell^{(t)})$ .

Assumptions on $\ell^{(t)}$	convex	$\alpha$ -exp-concave	uniformly $L$ -Lipschitz
Assumptions on $S$	convex	convex	$v$ -uniformly fat
Method	Gradient descent [24]	Hedge [17]	Dual Averaging with strongly convex $f$ -divergence s.t. $f(x) = \mathcal{O}(x^{1+\epsilon})$ (Section 3)
Learning rates	$t^{-\frac{1}{2}}$	$\alpha$	$t^{-\frac{1}{2+n\epsilon}}$
$R^{(t)}/t$	$\mathcal{O}(t^{-\frac{1}{2}})$	$\mathcal{O}(t^{-1} \log t)$	$\mathcal{O}\left(t^{-\frac{1}{2+n\epsilon}}\right)$

Table 1: Regret upper bounds for different classes of losses.

When the feasible set  $S$  is finite, and the losses  $(\ell^{(t)})$  are uniformly bounded, the Hedge algorithm [9], also known as the multiplicative weight updates [1] or the exponentiated gradient method [18], is known to achieve sublinear regret, and is easy to analyze and to implement. More general classes of algorithms with sublinear regret have been developed since. For example, the the online mirror descent algorithm [8], an extension of the mirror descent method due to Nemirovski and Yudin [20], is shown to have sublinear regret for any choice of strongly convex distance-generating function. Similarly, the dual averaging method [21] is shown to achieve sublinear regret for any choice of strongly convex regularizer. Remarkably, both of these families of algorithms include the Hedge algorithm as a special case.

When the set  $S$  is infinite, designing sublinear regret algorithms requires making additional assumptions on the class of loss functions  $(\ell^{(t)})$ , as well as the feasible set  $S$ . In [24], Zinkevitch considers an online problem on a convex  $S$ , for convex loss functions  $\ell^{(t)}$ . He shows that a simple gradient descent algorithm is guaranteed to have regret which grows as  $\mathcal{O}(\sqrt{t})$ . In [17], Hazan et al. also study the online learning problem on convex  $S$ , and show that for some classes of loss functions, one can achieve logarithmic regret, i.e.  $R^{(t)} = \mathcal{O}(\log t)$ . In particular, they show that logarithmic regret is achieved by the Newton method when the losses are  $\alpha$ -strongly convex, and by the Hedge algorithm when the losses are  $\alpha$ -exp concave (uniformly in  $t$ ).

In this article, we design sublinear regret algorithms under mild assumptions on the feasible set and the sequence of losses. In particular, we only assume that the losses are Lipschitz-continuous, and relax the convexity assumption on the set  $S$ . Our main result is summarized in Table 1, together with regret bounds for other classes of loss functions. We show that one can formulate the online learning problem as an optimization problem over a convex subset of  $L^2(S)$ , allowing us to use results from (infinite dimensional) convex analysis. By applying the dual averaging method of Nesterov to  $L^2(S)$  we prove, in Section 2, a general regret bound which holds for any choice of regularizer. In Section 3, we consider a particular class of regularizers, which can be expressed as Csiszár divergences of  $\omega$ -potentials, and we derive sufficient conditions on the potential to (i) make the dual averaging solution efficiently computable, and (ii) to guarantee that the regret grows sublinearly on any sequence of uniformly Lipschitz losses. This results in a general class of algorithms which are efficient to implement and which have sublinear regret guarantees under mild assumptions on the feasible set and the class of losses. In Section 5, we give concluding remarks, connections with related problems, and directions for future work.

## 2 Dual averaging on $L^2(S)$

We start by applying Nesterov’s dual averaging method [21] to our sequential decision problem viewed as an online optimization problem on a convex subset of  $L^2(S)$ , and derive a general regret bound for this algorithm.

### 2.1 Dual Averaging on a Hilbert space

Consider a Hilbert space  $E$ , and a feasible set  $\mathcal{X} \subset E$ , assumed to be closed and convex, and let  $\|\cdot\|$  be a reference norm on  $E$  (not necessarily the norm induced by the inner product).

Let  $\psi : \mathcal{X} \rightarrow \mathbb{R}_+$  be proper, continuous, and Fréchet-differentiable on the interior of  $\mathcal{X}$ , denoted by  $\mathring{\mathcal{X}}$ . The Bregman divergence associated to  $\psi$  is defined as follows:

$$D_\psi : \mathcal{X} \times \mathring{\mathcal{X}} \rightarrow \mathbb{R}_+$$

$$(x, y) \mapsto D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla \psi(y), x - y \rangle$$

The function  $\psi$  is said to be  $\ell_\psi$ -strongly convex with respect to a reference norm  $\|\cdot\|$  for all  $x, y \in \mathcal{X} \times \mathring{\mathcal{X}}$ ,

$$D_\psi(x, y) \geq \frac{\ell_\psi}{2} \|x - y\|^2$$

It is  $L_\psi$ -smooth with respect to  $\|\cdot\|$  if for all  $x, y \in \mathcal{X} \times \mathring{\mathcal{X}}$ ,

$$D_\psi(x, y) \leq \frac{L_\psi}{2} \|x - y\|^2$$

As we describe below, strong convexity and smoothness are dual properties. Define the Fenchel-Legendre conjugate of  $\psi$

$$\psi^*(y) = - \inf_{x \in \mathcal{X}} \psi(x) - \langle y, x \rangle$$

Note that the minimum is attained and the minimizer is unique since  $\psi$  is strongly convex and  $\mathcal{X}$  is closed and convex (Theorem 11.9 in [3]). The gradient of  $\psi^*$  is

$$\nabla \psi^*(y) = \arg \min_{x \in \mathcal{X}} \psi(x) - \langle y, x \rangle$$

which we will refer to as the Bregman projection onto  $\mathcal{X}$ , since it can be written as

$$\nabla \psi^*(y) = \arg \min_{x \in \mathcal{X}} \psi(x) - \langle \nabla \psi(\nabla \psi^{-1}(y)), x \rangle = \arg \min_{x \in \mathcal{X}} D_\psi(x, \nabla \psi^{-1}(y))$$

**Proposition 1.** *If  $\psi$  is  $\ell_\psi$ -strongly convex with respect to  $\|\cdot\|$ , then  $\psi^*$  is  $\frac{1}{\ell_\psi}$ -smooth with respect to the dual norm  $\|\cdot\|_*$ .*

Proposition 1 is an extension of Theorem 18.15 in [3] to general norms, the proof is provided in the Appendix.

Given a sequence  $(\ell^{(t)})$  of linear functionals in the dual space  $E^*$ , the method projects, at each step, the cumulative dual vector  $-L^{(t)} = -\sum_{\tau=1}^t \ell^{(\tau)}$ , scaled by a step size  $\eta_{t+1}$ , onto the feasible set, using the Bregman projection  $\nabla \psi^*$ . This is summarized in Algorithm 2. Without loss of generality, we will assume that  $\inf_{x \in \mathcal{X}} \psi(x) = 0$ .

---

**Algorithm 2** Dual averaging method with input sequence  $(\ell^{(t)})$  and learning rates  $(\eta_t)$

---

- 1: **for**  $t \in \mathbb{N}$  **do**
- 2:   Define  $L^{(t)} = \sum_{\tau=1}^t \ell^{(\tau)}$
- 3:   Update

$$x^{(t+1)} = \nabla \psi^*(-\eta_{t+1} L^{(t)}) = \arg \min_{x \in \mathcal{X}} \langle L^{(t)}, x \rangle + \frac{1}{\eta_{t+1}} \psi(x) \quad (1)$$

- 4: **end for**
-

## 2.2 Dual Averaging on $L^2(S)$

In particular, we consider the case where  $E = L^2(S)$ , the Lebesgue space of square integrable functions on  $S$ , endowed with the inner product  $\langle f, g \rangle = \int_S f(s)g(s)\lambda(ds)$ , where  $\lambda$  is the scaled Lebesgue measure such that  $\lambda(S) = 1$ . Let the feasible set  $\mathcal{X}$  be

$$\mathcal{X} := \{f \in L^2(S) : f \geq 0 \text{ a.e. and } \int_S f(s)\lambda(ds) = 1\}$$

Note that while  $\mathcal{X}$  is closed and convex, it is unbounded: if  $A$  is a measurable subset of  $S$ , then  $\frac{1}{\lambda(A)}1_A \in \mathcal{X}$ , and  $\|\frac{1}{\lambda(A)}1_A\|_2 = \frac{1}{\sqrt{\lambda(A)}}$ , which can be arbitrarily large.

An element  $f \in \mathcal{X}$  will be identified with the probability distribution on  $S$  with density  $f$ . The dual space is  $E^* = L^2(S)$ , and since  $S$  is compact,  $E^*$  contains, in particular, the set  $C^0(S)$  of continuous functions on  $S$ . Problem 1 can be viewed as follows: at each iteration  $t$ , the decision maker chooses an element of  $\mathcal{X}$ , then an element  $\ell^{(t)} \in C^0(S) \subset L^2(S)$  is revealed, and the decision maker incurs the expected loss  $\langle \ell^{(t)}, x^{(t)} \rangle$ . Next, we define the regret and provide a first bound on the regret of the dual averaging method.

**Definition 1.** Let  $(\ell^{(t)})$  be a sequence of elements of  $L^2(S)$ , and consider the dual averaging algorithm on this sequence, with learning rates  $(\eta_t)$ . The cumulative regret of the algorithm is defined as

$$R^{(t)} = \sup_{x \in \mathcal{X}} \sum_{\tau=1}^t \langle \ell^{(\tau)}, x^{(\tau)} - x \rangle = \sum_{\tau=1}^t \langle \ell^{(\tau)}, x^{(\tau)} \rangle - \inf_{x \in \mathcal{X}} \left\langle \sum_{\tau=1}^t \ell^{(\tau)}, x \right\rangle$$

The regret is said to be sublinear if  $\limsup_{t \rightarrow \infty} \frac{R^{(t)}}{t} \leq 0$ .

The regret compares the cumulative loss of the algorithm,  $\sum_{\tau=1}^t \langle \ell^{(\tau)}, x^{(\tau)} \rangle$ , to the best cumulative loss of any stationary distribution  $x$  (note that the infimum may not be attained).

**Lemma 1** (Dual Averaging Regret). Consider the dual averaging method with dual sequence  $(\ell^{(t)})$  and learning rates  $(\eta^{(t)})$ . Suppose that  $\psi$  is  $\ell_\psi$ -strongly convex w.r.t.  $\|\cdot\|$ , and that the losses are bounded in the dual norm, uniformly in  $t$ , i.e. there exists  $M > 0$  such that for all  $t$ ,  $\|\ell^{(t)}\|_* \leq M$ . Then for all  $t$  and all  $x \in \mathcal{X}$ ,

$$\sum_{\tau=1}^t \langle \ell^{(\tau)}, x^{(\tau)} - x \rangle \leq \frac{1}{\eta_t} \psi(x) + \frac{M^2}{2\ell_\psi} \sum_{\tau=1}^t \eta_\tau$$

*Proof.* Define the potential function

$$\xi(\eta, L) = -\frac{1}{\eta} \psi^*(-\eta L)$$

We first show the following inequality:

$$\langle x^{(t)}, \ell^{(t)} \rangle \leq \xi(\eta_t, L^{(t)}) - \xi(\eta_{t-1}, L^{(t-1)}) + \frac{\eta_t}{2\ell_\psi} \|\ell^{(t)}\|_*^2 \quad (2)$$

Since  $\psi$  is  $\ell_\psi$ -strongly convex w.r.t.  $\|\cdot\|$ , by Proposition 1,  $\psi^*$  is  $\frac{1}{\ell_\psi}$ -smooth w.r.t.  $\|\cdot\|_*$ , therefore  $D_{\psi^*}(-\eta_t L^{(t)}, -\eta_{t-1} L^{(t-1)}) \leq \frac{1}{2\ell_\psi} \|\eta_t L^{(t)} - \eta_{t-1} L^{(t-1)}\|_*^2$ , i.e.

$$\begin{aligned} \psi^*(-\eta_t L^{(t)}) - \psi^*(-\eta_{t-1} L^{(t-1)}) &\leq \left\langle \nabla \psi^*(-\eta_{t-1} L^{(t-1)}), -\eta_t L^{(t)} + \eta_{t-1} L^{(t-1)} \right\rangle + \frac{1}{2\ell_\psi} \|\eta_t L^{(t)} - \eta_{t-1} L^{(t-1)}\|_*^2 \\ &= -\eta_t \langle x^{(t)}, \ell^{(t)} \rangle + \frac{\eta_t^2}{2\ell_\psi} \|\ell^{(t)}\|_*^2 \end{aligned}$$

Thus

$$\left\langle x^{(t)}, \ell^{(t)} \right\rangle \leq \xi(\eta_t, L^{(t)}) - \xi(\eta_t, L^{(t-1)}) + \frac{\eta_t}{2\ell_\psi} \|\ell^{(t)}\|_*^2$$

It remains to show that  $\eta \mapsto \xi(\eta, G)$  is decreasing. Taking the derivative with respect to  $\eta$ ,

$$\begin{aligned} \partial_\eta \xi(\eta, L) &= \frac{1}{\eta^2} \psi^*(-\eta L) - \frac{1}{\eta} \langle \nabla \psi^*(-\eta L), -L \rangle \\ &= \frac{1}{\eta^2} (\psi^*(-\eta L) + \langle \nabla \psi^*(-\eta L), \eta L \rangle) \\ &\leq \frac{1}{\eta^2} \psi^*(0) && \text{by convexity of } \psi^* \\ &= -\frac{1}{\eta^2} \inf_{x \in \mathcal{X}} \psi(x) = 0 \end{aligned}$$

which proves inequality (2). Summing, and using the bound on  $\|\ell^{(t)}\|_*$ , we have

$$\sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle \leq \xi(\eta_t, L^{(t)}) - \xi(\eta_0, L^{(0)}) + \frac{M^2}{2\ell_\psi} \sum_{\tau=1}^t \eta_\tau$$

By definition of  $\xi$ , we have

$$\begin{aligned} \xi(\eta_0, L^{(0)}) &= \frac{1}{\eta_0} \psi^*(0) = -\frac{1}{\eta_0} \inf_{x \in \mathcal{X}} \psi(x) = 0 \\ \xi(\eta_t, L^{(t)}) &= \frac{1}{\eta_t} \inf_{x \in \mathcal{X}} \left\langle \eta_t L^{(t)}, x \right\rangle + \psi(x) \leq \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_t} \psi(x) \end{aligned}$$

Therefore

$$\sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle \leq \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_t} \psi(x) + \frac{M^2}{2\ell_\psi} \sum_{\tau=1}^t \eta_\tau$$

which proves the claim.  $\square$

Note that the regularizer  $\psi$  can be unbounded on  $S$ . This is true for example for the entropy regularizer  $\psi(x) = \int_S x(s) \ln(x(s)) \lambda(ds)$ , which we will use in Section 4. Thus to obtain a useful (sublinear) bound on the regret, it may not suffice to take a supremum in the bound of Lemma 1. This motivates the following Theorem. In what follows, we will assume that the loss functions are Lipschitz, uniformly in time. Let  $s_t^* \in \arg \min_{s \in S} L^{(t)}(s)$  (since the loss functions are continuous and  $S$  is compact, the minimum is attained). Intuitively, if the losses are Lipschitz, then  $\inf_{x \in \mathcal{X}} \langle L^{(t)}, x \rangle$  is well approximated by the cumulative loss of distributions which concentrate their mass around  $s_t^*$ .

**Theorem 1** (Dual Averaging Regret for Lipschitz Losses). *Suppose that  $\ell^{(t)}$  is  $L$ -Lipschitz, and  $\|\ell^{(t)}\|_* \leq M$ , uniformly in  $t$ . Then the dual averaging method with learning rates  $(\eta_t)$  guarantees the following bound on the regret: For any positive sequence  $(d_t)$ ,*

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2\ell_\psi} \frac{\sum_{\tau=1}^t \eta_{\tau+1}}{t} + Ld_t + \frac{1}{t\eta_{t+1}} \inf_{x \in \mathcal{B}_t} \psi(x) \quad (3)$$

where  $\mathcal{B}_t \subset \mathcal{X}$  denotes the set of Lebesgue-continuous densities supported on  $B(s_t^*, d_t)$ .

*Proof.* First, we observe that

$$R^{(t)} = \sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(t)} \right\rangle - \inf_{x \in \mathcal{X}} \langle x, \leq \rangle \sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle - L^{(t)}(s_t^*)$$

Since the losses are  $L$ -Lipschitz, we have  $\forall x \in \mathcal{B}_t$ ,

$$\left\langle \sum_{\tau=1}^t \ell^{(\tau)}, x \right\rangle = \int_{\mathcal{B}_t} \sum_{\tau=1}^t \ell^{(\tau)}(s) x(s) ds \leq \int_{\mathcal{B}_t} \sum_{\tau=1}^t (\ell^{(\tau)}(s_t^*) + L d_t) x(s) ds = L^{(t)}(s_t^*) + t L d_t$$

Thus, for all  $x \in \mathcal{B}_t$ ,

$$\begin{aligned} R^{(t)} &\leq \sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle - L^{(t)}(s_t^*) \\ &\leq \sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} - x \right\rangle + t L d_t \\ &\leq \frac{1}{\eta_{t+1}} \psi(x) + \frac{M^2}{2\ell_\psi} \sum_{\tau=1}^t \eta_{\tau+1} + t L d_t \end{aligned}$$

where the last inequality uses Lemma 1. We conclude by dividing by  $t$  and taking the infimum over  $x \in \mathcal{B}_t$ .  $\square$

We now have a general regret bound for the dual averaging method applied to Problem 1. In the next section, we further study the dual averaging algorithm with a particular family of regularizers, and we study their properties.

### 3 Dual Averaging with $\omega$ -potentials on Uniformly Fat Sets

We now study the dual averaging method when the regularizer is the  $f$ -divergence, or Csiszár divergence [11] of a particular class of potential functions. This definition is a generalization of [2] to our infinite dimensional Hilbert setting.

#### 3.1 Csiszár divergence induced by $\omega$ -potentials

**Definition 2.** Let  $\omega \leq 0$  and  $a \in (-\infty, +\infty]$ . An increasing  $C^1$  diffeomorphism  $\phi : (-\infty, a) \rightarrow (\omega, \infty)$ , is an  $\omega$ -potential if

$$\lim_{u \rightarrow -\infty} \phi(u) = \omega \quad \lim_{u \rightarrow a} \phi(u) = +\infty \quad \int_0^1 \phi^{-1}(u) du < \infty$$

We associate to  $\phi$  the function  $f_\phi$ , defined on  $(0, \infty)$ ,

$$f_\phi(x) = \int_1^x \phi^{-1}(u) du$$

which is, by definition, convex (since  $\phi^{-1}$  is increasing), and satisfies  $f_\phi(1) = 0^1$ . We also associate the  $f$ -divergence, defined on  $\mathcal{X}$  by

$$\psi_{f_\phi}(x) = \int_S f_\phi(x(s)) \lambda(ds)$$

By convexity of  $f$ , we have for all  $x \in \mathcal{X}$ ,  $\psi_{f_\phi}(x) \geq f_\phi(\int_S x(s) ds) = f_\phi(1) = 0$ .

**Example 1** (Euclidean projection). Perhaps the simplest instance of  $\omega$ -potential is the identity  $\phi(u) = u$ , for which  $\omega = -\infty$  and  $a = +\infty$ . In this case,  $f_\phi(x) = \frac{x^2-1}{2}$ , and the resulting Csiszár divergence is  $\psi_{f_\phi}(x) = \frac{1}{2} \int_S x(s)^2 \lambda(ds) - \frac{1}{2} = \frac{\|x\|_2^2 - 1}{2}$ . The dual averaging method then projects, at each iteration, in the  $L^2$  norm, the dual vector  $-\eta_{t+1} \sum_{\tau=1}^t \ell^{(\tau)}$  on the feasible set.

<sup>1</sup>Note that in the original definition of Audibert et al., the function  $f_\phi$  is taken to be the integral from  $\omega$  to  $\omega + x$ . Our definition corresponds to a translation of  $f_\phi$ , in order to have  $f_\phi(1) = 0$ , which guarantees that  $\psi_{f_\phi}(x) \geq 0$  on  $\mathcal{X}$ .

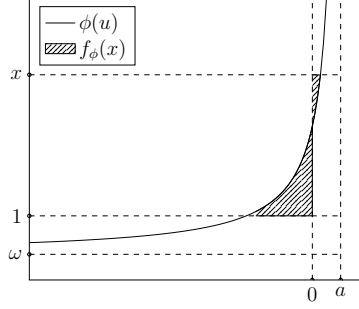


Figure 1: Illustration of an  $\omega$ -potential.

**Example 2.** More generally, if  $p > 1$ , then taking  $\phi(u) = \text{sign}(u)|u|^{\frac{1}{p-1}}$  is an  $\omega$ -potential with  $\omega = -\infty$  and  $a = +\infty$ . In this case,  $\phi^{-1}(u) = \text{sign}(u)|u|^{p-1}$  and  $f_\phi(x) = \frac{|x|^p}{p}$ , and the corresponding Csiszár divergence is  $\psi_{f_\phi}(x) = \frac{\|x\|_p^p - 1}{p}$ .

**Example 3** (Entropy projection). If we take  $\phi(u) = e^{u-1}$ , then the corresponding density function is  $f_\phi(x) = \int_1^x \phi^{-1}(u)du = \int_1^x (1 + \ln u)du = x \ln x$ , and the associated  $f_\phi$  divergence is the negative entropy

$$\psi_{f_\phi}(x) = \int_S x(s) \ln x(s) \lambda(ds)$$

See Section 4 for a generalization of the entropy divergence.

### 3.2 Strong convexity

In order for the bound of Theorem 1 to hold, we need the regularizer to be strongly convex. In this section, we give sufficient conditions on the potential for strong convexity of  $\psi_{f_\phi}$  with respect to  $p$  norms, defined as follows for  $p \geq 1$

$$\|x\|_p = \left( \int_S |x(s)|^p \lambda(ds) \right)^{\frac{1}{p}}.$$

**Theorem 2.** Let  $\phi$  be an  $\omega$ -potential, and suppose that there exists  $\alpha > 0$ ,  $z_0 \geq 0$  and  $r \in (0, 1]$  such that  $(\phi^{-1})'(z) \geq \frac{1}{\alpha(z+z_0)^r}$  for all  $z > 0$ . Then  $f_\phi$  is  $\ell$ -strongly convex w.r.t. the  $p$  norm, with  $\ell = \frac{1}{\alpha(1+z_0)^r}$  and  $p = \frac{2}{1+r}$ . That is, for all  $x, y \in \mathcal{X}$ ,

$$D_{f_\phi}(x, y) \geq \frac{1}{2\alpha(1+z_0)^r} \|x - y\|_{\frac{2}{1+r}}^2$$

*Proof.* By definition, the  $f$ -divergence associated to the potential  $\phi$  is differentiable at any  $x \in \mathcal{X}$  with  $x > 0$  a.e., and has gradient

$$\nabla \psi_{f_\phi}(x)(s) = \phi^{-1}(x(s))$$

Thus,

$$\begin{aligned} D_{f_\phi}(x, y) &= \psi(x) - \psi(y) - \langle \nabla \psi(y), x - y \rangle \\ &= \int_S f_\phi(x(s)) - f_\phi(y(s)) - \phi^{-1}(y(s))(x(s) - y(s)) \lambda(ds) \\ &= \int_S \int_{y(s)}^{x(s)} \phi^{-1}(u) du - \phi^{-1}(y(s))(x(s) - y(s)) \lambda(ds) \end{aligned}$$

and by a Taylor expansion of  $\phi^{-1}$ , there exists  $z \in [x, y]$  such that

$$\begin{aligned} \int_{y(s)}^{x(s)} \phi^{-1}(u) du - \phi^{-1}(y(s))(x(s) - y(s)) &= (\phi^{-1})'(z(s)) \frac{(x(s) - y(s))^2}{2} \\ &\geq \frac{1}{2} \frac{(x(s) - y(s))^2}{\alpha(z(s) + z_0)^r} \end{aligned} \quad \text{by assumption on } \phi$$

Now by the Cauchy-Schwartz inequality, if  $p, q \geq 1$  are conjugate, i.e.  $\frac{1}{p} + \frac{1}{q} = 1$ , then for any  $(a, b) \in \mathcal{H}^* \times \mathcal{H}$ ,  $\langle a, b \rangle \leq \|a\|_p \|b\|_q$ , and it follows that whenever  $\|b\|_q > 0$ ,  $\|a\|_p^p \geq \left( \frac{\langle a, b \rangle}{\|b\|_q} \right)^p$ . Applying this inequality with  $a = \left( \frac{(x-y)^2}{\alpha(z+z_0)^r} \right)^{\frac{1}{p}}$ ,  $b = (\alpha(z+z_0)^r)^{\frac{1}{p}}$ , we have

$$\int_S \frac{(x(s) - y(s))^2}{\alpha(z(s) + z_0)^r} \lambda(ds) \geq \left( \frac{\int_S (x(s) - y(s))^{\frac{2}{p}} \lambda(ds)}{\left( \int_S (\alpha(z(s) + z_0)^r)^{\frac{q}{p}} \lambda(ds) \right)^{\frac{1}{q}}} \right)^p = \frac{\|x - y\|_{\frac{2}{p}}^2}{\alpha \|z + z_0\|_{\frac{r}{p}}^r}$$

In particular, if we take  $p = 1 + r, q = 1 + \frac{1}{r}$ , then  $\frac{rq}{p} = 1$ , and  $\|z + z_0\|_{\frac{rq}{p}} = \|z + z_0\|_1 = 1 + z_0$  since  $z \in [x, y] \subset \mathcal{X}$  and  $z_0 \geq 0$ . Therefore

$$D_{f_\phi}(x, y) \geq \frac{1}{2} \int_S \frac{(x(s) - y(s))^2}{\alpha(z(s) + z_0)^r} \lambda(ds) \geq \frac{1}{2} \frac{\|x - y\|_{\frac{2}{1+r}}^2}{\alpha(1 + z_0)^r}$$

which concludes the proof.  $\square$

As a consequence of Theorem 2, we can show that Bregman divergences of Example 2,  $\psi_f = \frac{\|\cdot\|_p^p - 1}{p}$ ,  $p \in (1, 2]$ , are strongly convex w.r.t.  $\|\cdot\|_p$ .

**Corollary 1.** *Let  $p \in (1, 2)$ , and consider the  $\omega$ -potential  $\phi(u) = \text{sign}(u)|u|^{\frac{1}{p-1}}$ , and its corresponding Csiszár divergence  $\psi(x) = \frac{\|x\|_p^p - 1}{p}$ . Then  $\psi$  is  $(p-1)$ -strongly convex w.r.t.  $\|\cdot\|_{\frac{2}{3-p}}$ .*

*Proof.* We have  $\phi^{-1}(u) = \text{sign}(u)|u|^{p-1}$ , thus for  $z > 0$ ,

$$(\phi^{-1})'(z) = \frac{p-1}{z^{2-p}}$$

where  $2-p \in (0, 1)$  by assumption on  $p$ . Thus we can apply Theorem 2 with  $r = 2-p$ ,  $\alpha = \frac{1}{p-1}$  and  $z_0 = 0$ , which proves the claim.  $\square$

Note that the corollary also holds for  $p = 2$ , since one can explicitly compute the Bregman divergence: we have  $\psi(x) = \frac{\|x\|_2^2 - 1}{2}$ , thus

$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla \psi(y), x - y \rangle = \frac{1}{2}(\|x\|_2^2 - \|y\|_2^2 - 2\langle y, x - y \rangle) = \frac{1}{2}\|x - y\|_2^2$$

which is 1-strongly convex w.r.t.  $\|\cdot\|_2$ . Finally, we observe that a similar result is proved, in the finite-dimensional case, in [4], Lemma 8.1: for  $p \in (1, 2]$ ,  $\frac{1}{2}\|\cdot\|_p^2$  is  $(p-1)$ -strongly convex w.r.t.  $\|\cdot\|_p$ . Note that the result concerns  $\|\cdot\|_p^2$  while Corollary 1 concerns  $\|\cdot\|_p^p$ . The squared  $p$ -norm is not a Csiszár divergence induced by an  $\omega$ -potential in general (except for  $p = 2$ ). Using  $\|\cdot\|_p^p$  as a regularizer instead of  $\|\cdot\|_p^2$  allows us to benefit from the properties of  $\omega$ -potentials; in particular, the solution of the dual averaging iteration can be computed efficiently, as discussed in Section 3.3.

Next, we give another sufficient condition for strong convexity w.r.t.  $\|\cdot\|_1$ .



**Proposition 2.** Let  $\phi$  be a  $\omega$ -potential, and assume that  $\phi$  is a  $C^2$ -diffeomorphism. Consider the potential density  $f_\phi$  as in Definition 2. Let  $w = 1 + \frac{1}{3} \frac{f_\phi'''(1)}{f_\phi''(1)}$ . If  $f_\phi$  satisfies one of the following conditions  $\forall u > 0$ :

$$(f_\phi(u) - f_\phi'(1)(u-1))(1 + (1-w)(u-1)) \geq \frac{f_\phi''(1)}{2}(u-1)^2$$

$$\text{sgn}(u-1) \left( \frac{f_\phi'''(u)}{f_\phi''(u)} [1 + (1-w)(u-1)] + 3(1-w) \right) \geq 0,$$

then the  $f_\phi$ -divergence is strongly convex with respect to the total variation norm. More precisely, for all  $x, y \in \mathcal{X}$ ,

$$D_{f_\phi}(x, y) \geq \frac{1}{8\phi'(\phi^{-1}(1))} \|x - y\|_1^2$$

*Proof.* By Definition 2, if  $\phi$  is a  $C^2$  diffeomorphism, then  $f_\phi$  is three times differentiable on  $(0, \infty)$ , and for all  $x > 0$ ,  $f''(x) = (\phi^{-1})'(x) = \frac{1}{\phi'(\phi^{-1}(x))}$ , which is, by assumption on  $\phi$ , strictly positive. Thus by the generalized Pinsker inequality in [13] (Theorem 3 and Corollary 4) the associated  $f$ -divergence satisfies  $D_{f_\phi}(x, y) \geq \frac{f''(1)}{2} D_{TV}(x, y)^2$ , where  $D_{TV}$  is the total variation norm,  $D_{TV}(x, y) = \frac{1}{2} \|x - y\|_1$ . This concludes the proof.  $\square$

### 3.3 Solution of the Bregman projection with $\omega$ -potentials

We now characterize the solution of the dual averaging update, given by the Bregman projection in equation (1).

**Proposition 3.** Let  $\phi$  be an  $\omega$ -potential. Let  $L^{(t)} \in E^*$ , and consider the dual averaging iteration (1) in Algorithm 2, with the regularizer  $\psi$  taken to be the  $f_\phi$  divergence associated to  $\phi$ . Then the solution  $x^{(t+1)}$  is given by

$$x^{(t+1)}(s) = \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^*))_+$$

where  $y_+$  denotes the positive part of  $y$ , and  $\nu^*$  satisfies  $\int_S \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^*))_+ \lambda(ds) = 1$ .

*Proof.* Let  $K$  be the cone  $K = \{x \in L^2(S) : x \geq 0 \text{ a.e.}\}$ , and let

$$f(x) = \langle L^{(t)}, x \rangle + \frac{1}{\eta_{t+1}} \psi(x) + i_K(x)$$

where  $i_K$  is the indicator function of the cone  $K$ , i.e.  $i_K(s) = 0$  if  $s \in K$  and  $+\infty$  otherwise. The dual averaging iteration is equivalent to the following problem:

$$\begin{aligned} & \text{minimize}_{x \in L^2(S)} && f(x) \\ & \text{subject to} && \langle \mathbf{1}, x \rangle = 1 \end{aligned}$$

where  $\mathbf{1} : S \rightarrow \mathbb{R}$  is identically equal to 1. Using the fact that the subdifferential of the indicator  $i_K$  is the normal cone  $N_K$  (See for example Chapter 16 in [3]) given by  $\forall x \in K$

$$\partial i_K(x) = N_K(x) = \left\{ g \in L^2(S) : \sup_{y \in K} \langle g, y - x \rangle \leq 0 \right\},$$

the subdifferential of the objective function is

$$\partial f(x) = L^{(t)} + \frac{1}{\eta_{t+1}} \phi^{-1}(x) + N_K(x)$$

First, we show that, for all  $x$  and all  $g \in N_K(x)$ ,  $g \leq 0$  and  $gx = 0$  almost everywhere. Indeed, fixing  $x \in K$ , we have  $\langle g, y - x \rangle \leq 0$  for all  $y \in K$ . In particular,  $x + 1_{g>0} \in K$ , thus

$$0 \geq \langle g, 1_{g>0} \rangle = \int_S g_+(s) \lambda(ds)$$

which proves that  $g_+ = 0$  a.e.. Furthermore, taking  $y = \frac{1}{2}x \in K$ , we have

$$0 \geq \langle g, y - x \rangle = -\frac{1}{2} \langle g, x \rangle = \frac{1}{2} \int_S |g(s)| x(s) ds$$

which implies that  $|g|x = 0$  a.e., which proves the claim.

Now, consider the Lagrangian  $\mathcal{L} : E \times \mathbb{R} \rightarrow \mathbb{R}$

$$\mathcal{L}(x, \nu) = \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_{t+1}} \psi(x) + i_K(x) + \nu(\langle \mathbf{1}, x \rangle - 1)$$

Then  $(x^*, \nu^*)$  is an optimal pair only if

$$\begin{aligned} 0 &\in L^{(t)} + \frac{1}{\eta_{t+1}} \phi^{-1}(x^*) + N_K(x^*) + \nu \mathbf{1} \\ \langle \mathbf{1}, x^* \rangle &= 1 \end{aligned}$$

see for example Section 19.3 in [3]. We can rewrite the stationarity condition as follows:

$$\exists g^* \in N_K(x^*) \text{ such that } L^{(t)} + \frac{1}{\eta_{t+1}} \phi^{-1}(x^*) + \nu \mathbf{1} + g^* = 0.$$

Therefore,

$$\begin{aligned} x^*(s) &= \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^* + g^*(s))) \text{ a.e.} \\ g^* &\in N_K(x^*) \\ \langle \mathbf{1}, x^* \rangle &= 1 \end{aligned}$$

In particular, let  $\text{support}(x^*) = \{s : x^*(s) > 0\}$  (a measurable set). By the complementary slackness condition,  $g^* x^* = 0$  a.e., therefore  $g^* = 0$  a.e. on  $\text{support}(x^*)$ . And for a.e.  $s \notin \text{support}(x^*)$ , we have

$$\phi(-\eta_{t+1}(L^{(t)}(s) + \nu^*)) \leq \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^* + g^*(s))) = 0$$

since  $\phi$  is increasing and  $g^* \leq 0$  a.e.. Therefore the optimality conditions become

$$\begin{aligned} x^*(s) &= \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^*))_+ \\ \int_S \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^*))_+ ds &= 1 \end{aligned}$$

which proves the claim.  $\square$

Proposition 3 shows that the solution of the Bregman projection 1 is entirely determined by the dual variable  $\nu^*$ , therefore computing the solution reduces to computing the optimal  $\nu^*$ . Furthermore, we observe that the function  $\nu \mapsto \int_S \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^*))_+ \lambda(ds)$  is increasing, by assumption on  $\phi$ , therefore one can compute  $\nu^*$  (to arbitrary precision) using a simple bisection method. Note that in general, the solution  $x^{(t+1)}$  may not be supported everywhere on  $S$ , unless  $\omega = 0$ , in which case  $\phi$  is by definition, strictly positive.

### 3.4 Regret analysis

Next, we show that under the appropriate assumptions on the feasible set  $S$ , and the asymptotic behavior of the  $\omega$ -potential, it is possible to achieve sublinear regret with dual averaging. First, we focus our attention on sets which are uniformly fat (using the definition of [19]), a generalization of convexity.

**Definition 3.** Consider a subset  $S \subset \mathbb{R}^n$ .  $S$  is said to be  $v$ -uniformly fat if there exists  $v > 0$  such that for all  $s \in S$ , there exists a convex  $K_s \subset S$  such that  $s \in K_s$  and  $\lambda(K_s) \geq v$ .

In particular, if  $S$  is convex, it is 1-uniformly fat.

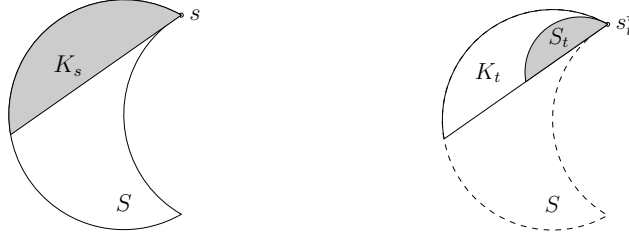


Figure 2: Illustration of uniform fatness (left) and the homothetic transformation used in the proof of Proposition 4 (right).

Intuitively, the uniform fatness condition guarantees that there is sufficient volume around any point of  $S$ , so that the solution  $x^{(t)}$  of the Bregman projection assigns enough probability mass around the optimum. In particular, uniform fatness excludes isolated points.

The next Proposition gives a regret bound on uniformly fat sets. For a subset  $C \subset \mathbb{R}^n$ , we denote the diameter of  $C$  by  $D(C) = \sup_{s, s' \in C} \|s - s'\|$ .

**Proposition 4.** Suppose that  $S$  is  $v$ -uniformly fat. Let  $\phi$  be an  $\omega$ -potential, and take the regularizer to be the Csiszár divergence  $f_\phi$ . Suppose that  $f_\phi$  is  $\ell_\psi$ -strongly convex w.r.t.  $\|\cdot\|$ , and that  $\|\ell^{(t)}\|_*$  is bounded by  $M$ . Let  $(d_t)$  be a sequence of positive numbers. Then for all  $t$ ,

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2\ell_\psi} \frac{\sum_{\tau=1}^t \eta_{\tau+1}}{t} + LD(S)d_t + \frac{1}{t\eta_{t+1}} v d_t^n f_\phi\left(\frac{1}{v d_t^n}\right) \quad (4)$$

*Proof.* Let  $s_t^* \in \arg \min_{s \in S} L^{(t)}(s)$ . Since  $S$  is  $v$ -uniformly fat, there exists a convex subset  $K_t \subset S$ , containing  $s_t^*$ , such that  $\lambda(K_t) \geq v$ . Following the argument in [19] and [17], consider the homothetic transformation of  $K_t$ , given by

$$S_t = \{d_t(s - s_t^*), s \in S\}$$

The diameter and mass of  $S_t$  satisfy

$$\begin{aligned} D(S_t) &= d_t D(K_t) \leq d_t D(S) \\ \lambda(S_t) &= d_t^n \lambda(K_t) \geq v d_t^n \end{aligned}$$

Let  $u_{S_t}$  be the uniform distribution over  $S_t$ . Since  $u_{S_t}$  has support in  $B(s_t^*, d_t D(S))$ , we have, by Theorem 1,

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2\ell_\psi} \frac{\sum_{\tau=1}^t \eta_{\tau+1}}{t} + LD(S)d_t + \frac{1}{t\eta_{t+1}} \psi(u_{S_t})$$

where, by definition of the  $f_\phi$  divergence,

$$\psi(u_{S_t}) = \int_{S_t} f_\phi(1/\lambda(S_t)) \lambda(ds) = f_\phi(1/\lambda(S_t)) \lambda(S_t)$$

To conclude, we observe that  $f_\phi$  is increasing on  $[1, \infty)$ . Indeed, by definition of an  $\omega$ -potential,  $f_\phi(x)$  has derivative  $\phi^{-1}(x)$ , and since  $\phi^{-1}$  is increasing, for all  $x \geq 1$ ,  $\phi^{-1}(x) \geq \frac{1}{1-\omega} \int_\omega^1 \phi^{-1}(u) du = 0$ . Therefore

$$\psi(u_{S_t}) = f_\phi(1/\lambda(S_t))\lambda(S_t) \geq v d_t^n f_\phi\left(\frac{1}{v d_t^n}\right)$$

which concludes the proof.  $\square$

Next, we show that if  $f_\phi$  has a bounded asymptotic growth rate, the regret grows sublinearly.

**Theorem 3.** *Suppose that  $S$  is  $v$ -uniformly fat. Let  $\phi$  be an  $\omega$ -potential such that  $f_\phi$  is  $\ell_\psi$ -strongly convex with respect to  $\|\cdot\|$ , and suppose that  $\ell^{(t)}$  is  $L$ -Lipschitz and  $\|\ell^{(t)}\|_* \leq M$  uniformly in  $t$ . Suppose that there exists  $\epsilon > 0$  and  $C > 0$  such that*

$$f_\phi(x) \leq Cx^{1+\epsilon}$$

*for all  $x \geq 1$ . Then the dual-averaging method with  $f_\phi$ -divergence and learning rates  $\eta_t = \theta t^{-\alpha}$  satisfies the following bound on the per-round regret*

$$\begin{aligned} \frac{R^{(t)}}{t} &\leq \frac{M^2\theta}{2\ell_\psi(1-\alpha)}t^{-\alpha} + (LD(S) + Cv^\epsilon)t^{-\frac{1-\alpha}{1+n\epsilon}} \\ &= \mathcal{O}\left(t^{-\alpha} + t^{-\frac{1-\alpha}{1+n\epsilon}}\right) \end{aligned}$$

*The rate is optimal for  $\alpha = \frac{1}{2+n\epsilon}$  for which*

$$\frac{R^{(t)}}{t} = \mathcal{O}\left(t^{-\frac{1}{2+n\epsilon}}\right)$$

*Proof.* Let  $(d_t)$  be a decreasing sequence which converges to 0. By Proposition 4, the regret of the dual averaging method is bounded by

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2\ell_\psi} \frac{\sum_{\tau=1}^t \eta_\tau}{t} + LD(S)d_t + \frac{1}{t\eta_t} v d_t^n f_\phi\left(\frac{1}{v d_t^n}\right)$$

where we can bound  $\sum_{\tau=1}^t \eta_\tau \leq \theta \sum_{\tau=1}^t \int_{\tau-1}^\tau u^{-\alpha} du = \theta \int_0^t u^{-\alpha} du = \frac{\theta}{1-\alpha} t^{1-\alpha}$ . By assumption on  $f_\phi$ , we have

$$v d_t^n f_\phi\left(\frac{1}{v d_t^n}\right) \leq C v^{-\epsilon} d_t^{-n\epsilon}$$

Combining these bounds, we have

$$\begin{aligned} \frac{R^{(t)}}{t} &\leq \frac{M^2\theta}{2\ell_\psi(1-\alpha)}t^{-\alpha} + LD(S)d_t + C v^\epsilon t^{\alpha-1} d_t^{-n\epsilon} \\ &= \mathcal{O}\left(t^{-\alpha} + d_t + t^{\alpha-1} d_t^{-n\epsilon}\right) \end{aligned}$$

Taking  $d_t = t^{-\frac{1-\alpha}{1+n\epsilon}}$ , the bound becomes  $\frac{R^{(t)}}{t} \leq \frac{M^2\theta}{2\ell_\psi(1-\alpha)}t^{-\alpha} + (LD(S) + C v^{-\epsilon})t^{-\frac{1-\alpha}{1+n\epsilon}}$ , which proves the claim.  $\square$

One can formulate similar regret bounds under different assumptions on the asymptotic behavior of  $f_\phi$ . For example, one can show the following extension, proved in the Appendix.

**Theorem 3.1.** *Under the assumptions of Theorem 3, suppose that there exists  $\epsilon > 0$ ,  $\nu \geq 0$  such that*

$$f_\phi(x) = \mathcal{O}\left(x^{1+\epsilon}(\ln x)^\nu\right)$$

*and the learning rates are taken to be  $\eta_t = \Theta\left((\ln t)^{\frac{\nu}{2}} t^{-\alpha}\right)$ ,  $\alpha = \frac{1}{2+n\epsilon}$  then*

$$\frac{R^{(t)}}{t} = \mathcal{O}\left((\ln t)^{\frac{\nu}{2}} t^{-\frac{1}{2+n\epsilon}}\right).$$

To conclude this Section, we observe that while the regret is defined with respect to elements of  $\mathcal{X}$  (Definition 1), it is equivalent, for uniformly fat sets, to the regret with respect to elements of  $S$ , in the following sense:

$$\begin{aligned} R^{(t)} &= \sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle - \inf_{x \in \mathcal{X}} \left\langle L^{(t)}, x \right\rangle \\ &= \sum_{\tau=1}^t \left\langle \ell^{(\tau)}, x^{(\tau)} \right\rangle - \min_{s \in S} L^{(t)}(s) \end{aligned} \quad (5)$$

*Proof.* Let  $s_t^* \in \arg \min_{s \in S} L^{(t)}(s)$ . Then it suffices to show that for all  $\epsilon > 0$ , there exists  $x \in \mathcal{X}$  such that

$$\left\langle \sum_{\tau=1}^t \ell^{(\tau)}, x \right\rangle \leq \sum_{\tau=1}^t \ell^{(\tau)}(s_t^*) + \epsilon$$

Fix  $\epsilon > 0$ . Since  $S$  is  $v$ -uniformly fat, there exists a convex set  $K_t \subset S$  containing  $s_t^*$ , with  $\lambda(K_t) > 0$ . Let  $\bar{S}_t$  be the homothetic transform of  $K_t$ , of center  $s_t^*$  and ratio  $d_t$ , as in the proof of Proposition 4. Then we have

$$\begin{aligned} D(\bar{S}_t) &= d_t D(K_t) \leq d_t D(S) \\ \lambda(\bar{S}_t) &= d_t^n \lambda(K_t) > 0 \end{aligned}$$

Now consider  $x = \frac{1}{\lambda(\bar{S}_t)} 1_{\bar{S}_t}$ . We have  $x \in \mathcal{X}$ , and since the  $\ell^{(\tau)}$  are uniformly  $L$ -Lipschitz,

$$\begin{aligned} \left\langle \sum_{\tau=1}^t \ell^{(\tau)}, x \right\rangle &= \sum_{\tau=1}^t \int_{\bar{S}_t} \frac{1}{\lambda(\bar{S}_t)} \ell^{(\tau)}(s) ds \\ &\leq \sum_{\tau=1}^t \int_{\bar{S}_t} \frac{1}{\lambda(\bar{S}_t)} (\ell^{(\tau)}(s_t^*) + L d_t D(S)) ds \\ &= t L d_t D(S) + \sum_{\tau=1}^t \ell^{(\tau)}(s_t^*) \end{aligned}$$

In particular, if we choose  $d_t = \frac{\epsilon}{t L D(S)}$ , we have  $\left\langle \sum_{\tau=1}^t \ell^{(\tau)}, x \right\rangle \leq \sum_{\tau=1}^t \ell^{(\tau)}(s_t^*) + \epsilon$ , which proves the claim.  $\square$

## 4 Entropy dual averaging

Let  $\omega \leq 0$ , and consider the  $\omega$ -potential

$$\begin{aligned} \phi : (-\infty, \infty) &\rightarrow (\omega, \infty) \\ u &\mapsto \phi(u) = e^{u-1} + \omega \end{aligned}$$

Its inverse is  $\phi^{-1}(u) = 1 + \ln(u - \omega)$ . In particular, we have  $\int_0^1 \phi^{-1}(u) du = (u - \omega) \ln(u - \omega) \big|_0^1 < \infty$ . The resulting density function is simply

$$f_\phi(x) = \int_1^x \phi^{-1}(u) du = (x - \omega) \ln(x - \omega) - (1 - \omega) \ln(1 - \omega)$$

and the associated  $f_\phi$  divergence is the following generalized negative entropy

$$\begin{aligned} \psi_{f_\phi}(x) &= \int_S (x(s) - \omega) \ln(x(s) - \omega) \lambda(ds) - (1 - \omega) \ln(1 - \omega) \\ &= H(x - \omega) - H(\mathbf{1} - \omega) \end{aligned}$$

where  $H(x) = \int_S x(s) \ln x(s) \lambda(ds)$ . By applying Proposition 3, we can derive the explicit solution of the dual averaging update (1).

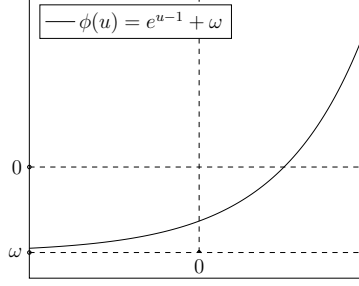


Figure 3: Illustration of the entropy  $\omega$ -potential  $\phi(u) = e^{u-1} + \omega$ .

**Corollary 2.** *Consider the dual averaging method on the set of distributions  $\mathcal{X}$ , regularized by the negative entropy. The solution of update (1) is given by*

$$x^{(t+1)}(s) = \left( \frac{e^{-\eta_{t+1} L^{(t)}(s)}}{Z^{(t)}} + \omega \right)_+$$

where  $Z^{(t)}$  is the appropriate normalization constant. In particular, when  $\omega = 0$ ,  $x^{(t+1)}(s) = \frac{e^{-\eta_{t+1} L^{(t)}(s)}}{Z^{(t)}}$  and  $Z^{(t)} = \int_S e^{-\eta_{t+1} L^{(t)}(s)} \lambda(ds)$ , we recover the Hedge algorithm [17].

By definition of the entropy potential function, we have  $(\phi^{-1})'(z) = \frac{1}{z-\omega}$ , therefore, the assumptions of Theorem 2 hold with  $\alpha = 1$ ,  $z_0 = -\omega$ ,  $r = 1$ , thus the negative entropy is  $\frac{1}{1-\omega}$ -strongly convex with respect to  $\|\cdot\|_1$ , and we can apply Proposition 4 and Theorem 3 to obtain a regret bound.

**Corollary 3.** *Suppose that  $S$  is  $v$ -uniformly fat, and that the loss functions are  $L$ -Lipschitz, uniformly in time, and that  $\|\ell^{(t)}\|_\infty \leq M$ . Then the dual averaging method on  $\mathcal{X}$ , regularized with the negative entropy, with learning rates  $\eta_t = \Theta\left(\sqrt{\frac{\ln t}{t}}\right)$ , has a sublinear regret such that  $\frac{R^{(t)}}{t} = \mathcal{O}\left(\sqrt{\frac{\ln t}{t}}\right)$ .*

*Proof.* By definition of the negative entropy, we have  $f_\phi(x) = (x-\omega) \ln(x-\omega) - (1-\omega) \ln(1-\omega) = \mathcal{O}(x \ln x)$ , and the results follows by Theorem 3.1.  $\square$

In fact, we can obtain a more explicit upper bound on the regret. By Proposition 4, we have

$$\frac{R^{(t)}}{t} \leq \frac{(1-\omega)M^2}{2} \frac{\sum_{\tau=1}^t \eta_{\tau+1}}{t} + LD(S)d_t + \frac{1}{t\eta_{t+1}} v d_t^n f_\phi\left(\frac{1}{v d_t^n}\right)$$

and we bound  $f_\phi\left(\frac{1}{v d_t^n}\right)$ . First, we have

$$\begin{aligned} f_\phi(x) &= (x-\omega) \ln(x-\omega) - (1-\omega) \ln(1-\omega) \\ &\leq (x-\omega) \ln(x-\omega) \\ &\leq c_\omega x \ln x + d_\omega \end{aligned} \quad \text{for } x \geq 1$$

where  $c_\omega = 1 + \ln(1-\omega)$  and  $d_\omega = (1-\omega) \ln(1-\omega)$ . To prove the last inequality, let  $\delta(x) = (x-\omega) \ln(x-\omega) - c_\omega x \ln x + d_\omega$ . Then  $\delta(1) = 0$  and for all  $x \geq 1$ ,

$$\delta'(x) = \ln(x-\omega) + 1 - c_\omega(1 + \ln x) = \ln \frac{x-\omega}{x^{c_\omega}(1-\omega)} \leq \ln \frac{x(1-\omega)}{x^{c_\omega}(1-\omega)} = \ln x^{1-c_\omega} \leq 0$$

Thus, whenever  $d_t \leq 1$ , we have  $\frac{1}{vd_t^n} \geq 1$ , and

$$vd_t^n f_\phi\left(\frac{1}{vd_t^n}\right) \leq c_\omega \ln \frac{1}{vd_t^n} + d_\omega vd_t^n$$

and taking  $d_t = 1/t$ , we have

$$\frac{R^{(t)}}{t} \leq \frac{(1-\omega)M^2}{2} \frac{\sum_{\tau=1}^t \eta_{\tau+1}}{t} + \frac{LD(S)}{t} + \frac{c_\omega(n \ln t - \ln v) + d_\omega v/t^n}{t \eta_{t+1}}$$

In particular, when  $\omega = 0$  (i.e. for the Hedge algorithm),  $c_\omega = 1$  and  $d_\omega = 0$ , and the bound simplifies to

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2} \frac{\sum_{\tau=1}^t \eta_{\tau+1}}{t} + \frac{LD(S)}{t} + \frac{n \ln t - \ln v}{t \eta_{t+1}}.$$

and we recover the bound on the Hedge regret obtained in [19].

## 5 Concluding Remarks

We studied a sequential problem in which a decision maker chooses, at each iteration, a distribution  $x^{(t)}$  over a compact set  $S$ , then observes a loss  $\ell^{(t)}$  from the class of  $L$ -Lipschitz continuous functions on  $S$ . Viewing the problem as an online convex problem over  $L^2(S)$ , we applied the dual averaging method and derived a general regret bound. Then we studied dual averaging with Csiszár divergences induced by  $\omega$ -potentials, and showed that for this class, the Bregman projection (1) can be computed efficiently, assuming one can efficiently evaluate integrals on  $S$  (e.g. by using a MCMC method). We then provided sufficient conditions on the asymptotic behavior of the potential to guarantee (i) strong convexity of the Csiszár divergence (Theorem 2), and (ii) a sublinear regret (Theorem 3). These sufficient conditions provide guidance in the design and analysis of dual averaging methods, which we illustrated with one particular family of  $\omega$ -potentials, using entropy regularizers.

Another approach for learning on a continuum consists in applying a discrete learning algorithm on a finite cover of  $S$ . Since the loss functions are  $L$ -Lipschitz, the additional regret incurred due to learning on the cover is at most  $L$  times the diameter of each element of the cover. Therefore, in order to have asymptotically sublinear regret, one would need to refine the cover as the number of iterations grows. Our method does not require explicitly computing a cover, since it samples directly from a distribution defined on  $S$ . It has the potential of being more computationally tractable (since one does not need to compute and refine a cover) but ultimately, its complexity is that of sampling from the sequence of distributions  $(x^{(t)})$ , thus studying the computational complexity of the proposed dual averaging method requires making additional assumptions on the family of loss functions and the feasible set  $S$ .

A related problem is bandit learning, in which the decision maker plays, at iteration  $t$ , an action  $s^{(t)}$  drawn from the distribution  $x^{(t)}$ , then only observes the loss of that action,  $\ell^{(t)}(s^{(t)})$ , as opposed to the full loss function  $\ell^{(t)}$ . This problem is studied for example in [7], for Lipschitz losses, and when the feasible set  $S$  is given by an explicit hierarchical formulation. We are currently investigating extensions of this work to the bandit setting.

## References

- [1] Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.
- [2] Jean-Yves Audibert, Sébastien Bubeck, and Gábor Lugosi. Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1):31–45, 2014.
- [3] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer, 2011.
- [4] Aharon Ben-Tal, Tamar Margalit, and Arkadi Nemirovski. The ordered subsets mirror descent optimization method with applications to tomography. *SIAM J. on Optimization*, 12(1):79–108, January 2001.
- [5] David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.
- [6] Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999.
- [7] S. Bubeck, R. Munos, G. Stoltz, and C. Szepesvari. X-armed bandits. *Journal of Machine Learning Research (JMLR)*, 12(12):1587–1627, 2011.
- [8] Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012.
- [9] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.
- [10] Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.
- [11] I. Csisz’ar. Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, 2:299–318, 1967.
- [12] Ofer Dekel, Ran Gilad-Bachrach, Ohad Shamir, and Lin Xiao. Optimal distributed online prediction. In *Proceedings of the 28th International Conference on Machine Learning (ICML)*, June 2011.
- [13] G.L. Gilardoni. On pinsker’s and vajda’s type inequalities for csiszár’s f-divergences. *Information Theory, IEEE Transactions on*, 56(11):5377–5386, Nov 2010.
- [14] James Hannan. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957.
- [15] Sergiu Hart and Andreu Mas-Colell. A general class of adaptive strategies. *Journal of Economic Theory*, 98(1):26 – 54, 2001.
- [16] Sergiu Hart and Andreu Mas-Colell. Regret-based continuous-time dynamics. *Games and Economic Behavior*, 45(2):375 – 394, 2003. Special Issue in Honor of Robert W. Rosenthal.
- [17] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.
- [18] Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132(1):1 – 63, 1997.
- [19] Walid Krichene, Maximilian Balandat, Claire Tomlin, and Alexandre Bayen. The hedge algorithm on a continuum. In *International Conference on Machine Learning (ICML)*, 2015.



- [20] A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization*. Wiley-Interscience series in discrete mathematics. Wiley, 1983.
- [21] Yurii Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259, 2009.
- [22] Noah D. Stein, Pablo A. Parrilo, and Asuman Ozdaglar. Correlated equilibria in continuous games: Characterization and computation. *Games and Economic Behavior*, 71(2):436 – 455, 2011.
- [23] Gilles Stoltz and Gábor Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187 – 208, 2007.
- [24] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.

## A Omitted proofs

### A.1 Proof of Proposition 1

**Proposition 1** If  $\psi$  is  $\ell_\psi$ -strongly convex with respect to  $\|\cdot\|$ , then  $\psi^*$  is  $\frac{1}{\ell_\psi}$ -smooth with respect to the dual norm  $\|\cdot\|_*$ .

*Proof.* Let  $y_1, y_2 \in E^*$ , and  $x_i = \nabla\psi^*(y_i)$ . Since  $x_i$  is the minimizer of the convex function  $x \mapsto \psi(x) - \langle y_i, x \rangle$ , we have, by first-order optimality,

$$\langle \nabla\psi(x_i) - y_i, x - x_i \rangle \geq 0 \quad \forall x \in \mathcal{X}$$

In particular, we have

$$\begin{aligned} \langle \nabla\psi(x_1) - y_1, x_2 - x_1 \rangle &\geq 0 \\ \langle \nabla\psi(x_2) - y_2, x_1 - x_2 \rangle &\geq 0 \end{aligned}$$

and summing both inequalities,

$$\langle y_2 - y_1, x_2 - x_1 \rangle \geq \langle \nabla\psi(x_2) - \nabla\psi(x_1), x_2 - x_1 \rangle$$

By definition of the Bregman divergence, we have

$$\begin{aligned} \langle y_2 - y_1, x_2 - x_1 \rangle &= \langle \nabla\psi(x_2) - \nabla\psi(x_1), x_2 - x_1 \rangle \\ &= D_\psi(x_1, x_2) + D_\psi(x_2, x_1) \\ &\geq \ell_\psi \|x_2 - x_1\|^2 \end{aligned}$$

and by definition of the dual norm, we have  $\langle y_2 - y_1, x_2 - x_1 \rangle \leq \|y_2 - y_1\|_* \|x_2 - x_1\|$ , since  $\|y\|^* = \sup_{\|x\| \leq 1} \langle x, y \rangle \geq \left\langle \frac{x}{\|x\|}, y \right\rangle$ . Therefore,

$$\|y_2 - y_1\|_* \|x_2 - x_1\| \geq \ell_\psi \|x_2 - x_1\|^2$$

rearranging, we have  $\|x_2 - x_1\| \leq \frac{1}{\ell_\psi} \|y_2 - y_1\|_*$ , i.e.

$$\|\nabla\psi^*(y_2) - \nabla\psi^*(y_1)\| \leq \frac{1}{\ell_\psi} \|y_2 - y_1\|_* \tag{6}$$

Now by definition of the Bregman divergence, we have

$$\begin{aligned}
D_{\psi^*}(x, y) &= \psi^*(x) - \psi^*(y) - \langle \nabla \psi^*(y), x - y \rangle \\
&= \int_0^1 \langle \nabla \psi^*(y + t(x - y)) - \nabla \psi^*(y), x - y \rangle dt \\
&\leq \|y - x\|_* \int_0^1 \|\nabla \psi^*(y + t(x - y)) - \nabla \psi^*(y)\| dt \\
&\leq \|y - x\|_* \int_0^1 \frac{1}{\ell_\psi} \|y + t(x - y) - y\|_* dt && \text{by (6)} \\
&\leq \frac{1}{\ell_\psi} \|x - y\|_*^2 \int_0^1 t dt \\
&= \frac{1}{\ell_\psi} \|x - y\|_*^2 \frac{1}{2}
\end{aligned}$$

□

## A.2 Proof of Theorem 3.1

**Theorem 3.1** Under the assumptions of Theorem 3, suppose that there exists  $\epsilon > 0$ ,  $\nu \geq 0$  such that

$$f_\phi(x) = \mathcal{O}(x^{1+\epsilon}(\ln x)^\nu)$$

and the learning rates are taken to be  $\eta_t = \Theta((\ln t)^{\frac{\nu}{2}} t^{-\alpha})$ ,  $\alpha = \frac{1}{2+n\epsilon}$  then

$$\frac{R^{(t)}}{t} = \mathcal{O}\left((\ln t)^{\frac{\nu}{2}} t^{-\frac{1}{2+n\epsilon}}\right).$$

*Proof.* The proof is similar to that of Theorem 3. Let  $(d_t)$  be a decreasing sequence which converges to 0. By Proposition 4,

$$\frac{R^{(t)}}{t} \leq \frac{M^2}{2\ell_\psi} \frac{\sum_{\tau=1}^t \eta_\tau}{t} + LD(S)d_t + \frac{1}{t\eta_t} v d_t^n f_\phi\left(\frac{1}{v d_t^n}\right)$$

where  $\sum_{\tau=1}^t \eta_\tau = \mathcal{O}((\ln t)^{\frac{\nu}{2}} t^{1-\alpha})$ ,  $\frac{1}{t\eta_t} = \mathcal{O}((\ln t)^{-\frac{\nu}{2}} t^{\alpha-1})$ , and by assumption on  $f_\phi$ ,

$$v d_t^n f_\phi\left(\frac{1}{v d_t^n}\right) \leq C v^{-\epsilon} d_t^{-n\epsilon} \left(\ln \frac{1}{v} + n \ln \frac{1}{d_t}\right)^\nu = \mathcal{O}\left(d_t^{-n\epsilon} \left(\ln \frac{1}{d_t}\right)^\nu\right)$$

Combining these bounds, we have

$$\frac{R^{(t)}}{t} = \mathcal{O}\left((\ln t)^{\frac{\nu}{2}} t^{-\alpha} + d_t + d_t^{-n\epsilon} t^{-(1-\alpha)} (\ln t)^{-\frac{\nu}{2}} \left(\ln \frac{1}{d_t}\right)^\nu\right)$$

Taking  $d_t = t^{-\frac{1-\alpha}{1+n\epsilon}}$ , the bound becomes

$$\frac{R^{(t)}}{t} = \mathcal{O}\left((\ln t)^{\frac{\nu}{2}} t^{-\alpha} + t^{-\frac{(1-\alpha)}{1+n\epsilon}} (\ln t)^{\frac{\nu}{2}}\right)$$

which proves the claim. □